B8

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 June 2001 (07.06.2001)

PCT

(10) International Publication Number
**WO 01/40896 A2**

(51) International Patent Classification[7]: G06F

(21) International Application Number: PCT/US00/33069

(22) International Filing Date: 6 December 2000 (06.12.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/454,991 6 December 1999 (06.12.1999) US

(71) Applicant: NATIONAL CENTER FOR GENOME RESOURCES [US/US]; 2935 Rodeo Park Drive East, Santa Fe, NM 87505 (US).

(72) Inventor: MENDES, Pedro; Apt. 512, 2500 Sawmill Road, Santa Fe, NM 87505 (US).

(74) Agents: ROBERTS, Jon, L. et al.; Roberts Abokhair & Mardula, LLC, Suite 1000, 11800 Sunrise Valley Drive, Reston, VA 20191 (US).

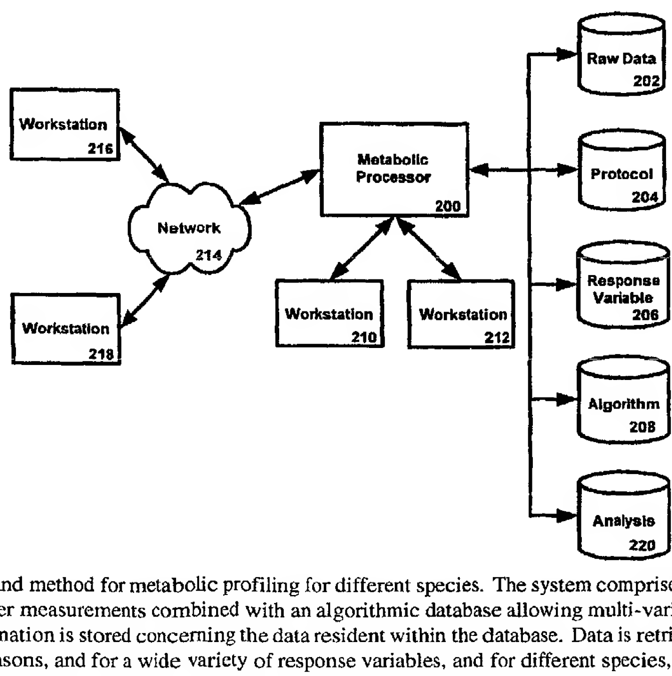
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**
— *Without international search report and to be republished upon receipt of that report.*

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: SYSTEM AND METHOD FOR METABOLIC PROFILING




(57) Abstract: A system and method for metabolic profiling for different species. The system comprises data from a wide variety of laboratory sensors and other measurements combined with an algorithmic database allowing multi-variate analysis of the data to be performed. Protocol information is stored concerning the data resident within the database. Data is retrieved throughout the growing season, across growing seasons, and for a wide variety of response variables, and for different species, thus allowing a wide variety of analysis across species and within species to take place. Data is stored concerning the workflow of any particular analysis so that scientists can build upon successful analyses to create further novel methods of analyzing metabolic data.

WO 01/40896 A2

**Title:** System and Method for Metabolic Profiling

**Field of the Invention:**

This invention relates generally to determining characteristics for healthy or beneficial agricultural products. More particularly, the present invention is a system and method for creating and analyzing metabolic profiles of plants for purpose of optimizing present and future agricultural products.

**Background of the Invention:**

In any plant species, genetic and chemical variation for a new growth cycle of the species can lead to a good or a poor harvest. Further, various diseases and environmental conditions can affect the yield of a particular crop sometimes with very beneficial and other times very disastrous results. For example, in the early seventies a corn blight manifested itself throughout the Midwestern part of the United States. This corn blight wiped out approximately 25% of the corn harvest yielding many billions of dollars of agricultural damage not to mention a ripple effect on prices for corn products as well as products that relied upon corn for feed purposes. The end result was billions of dollars of damage throughout the economy of the United States.

This led to a great deal of research on such things as environmental conditions, amount of fertilizer, and method of detection at an early stage when beneficial or detrimental circumstances exist with a particular crop. These various activities intended to focus on individual environmental factors as well as other man made factors such as amount of fertilizer applied and the like. The response variable for these individual studies has tended to be the amount of crop harvested. While this is certainly a useful "bottom line" figure, it does little to evaluate a trend in a crop to be productive or not. Yet there are many aspects throughout the growth cycle of plants that can be analyzed and correlated with good or bad harvest. For example, the near infrared spectrum of a crop and the absorption in the chlorophyll absorption band tend to change early when stress of one sort of another occurs in a given crop. This indicator can be detected early. Further,

various metabolic process occur during the growth cycle of a plant. The result of these metabolic processes are a series of "metabolites" which are present in the particular plant as a result of the various growth cycle functions. These metabolites will differ from plant to plant and will also differ during the course of the growth cycle of the plant. Using various technologies a wide variety of indicators of the metabolic activity of a plant species can be measured. For example, and without limitation, gas chromatography, high performance liquid chromatography, liquid chromatography mass spectrometry, ultra violet through the visible range of spectrophotometry both narrow band and wide band, thin layer chromatography, and infrared spectrometry in the near infrared midwave infrared and Far IR are all candidates technologies for measurement of the metabolights of various plant species.

In addition to these precise laboratory measurements, data can be recorded concerning weather, soil, plant physiology, geography, hydrology, genotype, gene expression, primary metabolites produced, processes metabolites can all be measured and stored for subsequent analysis.

While the above addresses issues of various plant varieties during the growing season there is also a significant post harvest issue that results when subsequent chemical processes, both natural and man made, continue after a plant species is harvested. For example, the ripening of fruits of all different kinds is a process that occurs frequently after the particular fruit in questions has been harvested. Knowing when to harvest, bananas, apples, peaches, and grapes to name but a few and, given the known time to market, can yield a better product on the shelves. The ripening process is a continuing chemical and metabolic process that continues in plants after they are harvested. The same holds true for various beverages that result from brewing and fermentation process such as beer and wine. Determining when the appropriate metabolic characteristics of a successful flavor exists can be critical to the generation of revenue and the building of the reputation of a particular organization for producing that particular product.

Further, each different type of agricultural crop has some similar measures, in the case of environmental conditions such as humidity, rain fall, days of rain in a given period, days of drought in a given period, and the like. In addition however, each crop

also has its own measurement characteristics and terms associated with the crop in question. Taste in a particular crop such as wine, fruit, and edible crops are very important. However, other types of crops, which are used for widely different purposes such as wood products, rely on totally different measurements. Thus different amounts of the same metabolites among species of plants do not necessarily symbolize a positive or negative trend in any of the species individually.

The monitoring of genetic expression also cannot be over emphasized. The fact that certain genes mutate in small or large ways also does not mean that the result in harvest will be bad. Indeed a gene mutation may led to a beneficial development in a particular crop. Once such gene mutations are uncovered and are determined to be beneficial, such genetic information, which results in enhanced traits of interest, maybe desired to be replicated to create entire crops comprising such mutated genes. Such crops could then be used for their normal commercial purposes or in the pharmaceutical field for example.

Given the above situation, wherein a variety of historical measurement variables exist and have been recorded and will be recorded in the future, where environmental information for a given crop also exists and can be measured in the future, where new laboratory techniques exist for measuring a variety of response variables which historically have no parallel, and given the fact that all of these different types of variables are in different quantities for different species of plants and indeed may exist for one species of plant and not for another, one has an analytical environment with unlimited and uncorrelated measurement combinations. Indeed, only some of these various response variables have only been used in a simplistic fashion to roughly correlate quality of crop with a single particular variable. Yet there is no comprehensive integrated way of analyzing all of these variables to determine how they co-vary and what the significance of that co-variance is.

What would be truly useful is a system and method that can be universally applied on a species by species basis to measure a wide variety of variables, determine how those variable co-vary with one an other and what the significance of that co-variance truly is. Such a system would utilize past historical measurement variables to provide appropriate

context for professionals operating in the particular agricultural arena in question, would allow for the input of new response variables as laboratory techniques become increasingly sophisticated, and would permit correlation analysis of these various types of environmental, laboratory, and historical types of measurement to determine in an efficacious way how best to manage and harvest a wide variety of agricultural crops and products.

What would further be useful is a transparent user friendly system which integrates data from various data sources which vary by data type and vary with species being analyzed. Such a system would also comprise a variety of analysis tools for multi-variate analysis, modeling, simulation, and visualization of resultant data. The system would also be able to take the results of successful research and store the steps used to obtain that successful research so that the scientific inferences and steps to achieve them are available for others to build upon.

**Summary of the Invention:**

It is therefore an objective of the present invention to record data, analyze a variety of data types, and predict the direction, either efficacious or not, for a particular crop during the growing season.

It is yet another objective of the present invention to analyze and predict, post harvest, the direction, efficacious or not, for agricultural products that are being stored.

It is a further objective of the present invention to be able to predict the direction, efficacious or not, for agricultural products that are the subject of post harvest processing.

It is yet another objective to create a system and method for storing, acquiring, and displaying all manner of metabolic data.

It is yet another objective of the present invention to store and correlate the experimental protocol used to obtain various types of metabolic data.

It is yet another objective of the present invention to store associated environmental and developmental data for a particular species.

It is a further objective of the present invention to be able to compare and analyze metabolic data for the same species under a variety of environmental conditions.

It is further objective of the present invention to be able to store and analyze metabolic data for the given species during various developmental stages in a growing season.

It is a further objective of the present invention to analyze data on the same species across various growing seasons at similar developmental stages.

It is yet another objective of the present invention to be able to analyze genetic mutations of the same species under the same environmental conditions.

It is a further objective of the present invention to analyze the metabolic profile variation of different species under the same environmental conditions.

It is yet another objective of the present invention to perform a multi-variate analysis on a wide variety of environmental, metabolic, and genetic information of a particular species to determine optimal growth conditions so that a given species will display an optimal set of traits.

It is yet another objective of the present invention to perform a multi-variate evaluation of metabolic, genetic, environmental conditions, to determine the optimal time to harvest the particular species in question.

It is yet another objective of the present invention to provide for the multi-variate analysis of metabolic, environmental, and genetic information of a particular species to characterize and optimize mutant strains for a particular species.

It is a further objective of the present invention to be able to browse and query data of various types representing various species in a user friendly fashion.

It is yet another objective of the present invention to be able to integrate different data regardless of database schemata, semantics, or syntax.

It is a further objective of the present invention to allow scientists to visualize information via a single system without having to download multiple tools from multiple locations and to convert files from one data type to another.

It is yet another objective of the present invention to allow scientists to have analysis algorithms tightly integrated with database resources so that multi-variate analysis tools are readily available.

It is yet another objective of the present invention to allow scientists to define workflows which can be executed repetitively on large batches of data from one growing season to another.

It is a further objective of the present invention to allow scientists to store discoveries or inferences of metabolic data so that they can be used to build upon for further research.

These and other objectives of the present invention will be apparent to those skilled in the art from a review of the specification that follows.

The present invention is a method and apparatus for multi-variate metabolic profiling that provides for a correlation between metabolic, genetic, and environmental factors that effect the growth of a particular species. The present invention allows a disciplined approach to the analysis and determination of optimal set of traits that characterize a successful crop. Conversely, the present invention allows for multi-variate analysis to determine when the particular crop is trending toward a detrimental harvest so that intervention can occur at an early stage thereby preventing economic hardship and reduced productivity.

The present invention comprises a database and processing capability for accepting a wide variety of input from field readings, environmental readings, and laboratory readings for a particular species and for analyzing those various readings in a multi-variate fashion to make deterministic analysis for the species in question.

For each variable being measured, raw data on the variable is stored together with the sampling protocol used to obtain the data. Further, data is stored based upon the species from which the data is collected. Thus for each species, raw data and a wide variety of field, environmental, and laboratory types will be stored together with the protocol used to collect each of the different types of data. This information is also stored according to the species about which the data is collected In this fashion, a multi dimensional data base is created.

Whole-organism metabolic profile data is also obtained and stored for a specific species. For example, and without limitation, gas chromatography - mass spectrometry (GC-MS), high performance liquid chromatography (HPLC), high performance liquid

chromatography-mass spectrometry (LCMS), thin layer chromatography (TLC), ultra violet and visible spectrophotometry (UV-VIS), short wavelength, mid wavelength, and long wavelength infrared (SWIR, MWIR, LWIR), raman spectrometry, and biosensor information are all collected and stored within the data base of the present invention. This information is collected for any given species desired, and is collected during various stages of the growing season, and over longer periods of time, on a season by season basis for the same variables during the same periods of respective growing seasons. As noted above the protocol for the collection of each type of this information is stored and associated with the data being stored.

A separate portion of the data base stores environmental information such as temperature, pressures, humidity, concentrations of exogenous (added) compounds such as fertilizer and other physical and chemical properties which may controlled by the crop owner or the experimentalist.

A library of various algorithmic approaches is also stored together with association of which multi variate analysis technique is best for a particular species, response variable, or other characteristic. Thus multi-variate analysis and/or principal axes factor analysis or other multi variate analysis algorithms are stored in an algorithmic data base and associated with the various species and response variables which are best used in conjunction with an algorithmic analysis.

The system further provides an automated method for data retrieval and analysis as well as numeric and visual display of data to optimize the human factors interaction with the very complex data base of the present invention.

Thus if, for example, a particular analyst wishes to review the trend for an orange crop the analyst would select the crop to determine the end product desired. For example, for vitamin production or orange juice production, the analyst would input the desired end product, and information from the multi-dimensional data base will be automatically retrieved, based upon those variables which are most predictive in nature, the appropriate analysis algorithm will be selected, the data will analyzed, and an appropriate output will be created for the analyst noting, among other things, the direction for the crop, (efficacious or not) whether intervention steps must be taken, what those steps should be,

and what the protocol would be for future analysis to monitor the crop in question. In addition, the system of the present invention has the ability to generate a sampling plan for a researcher or farm owner who wishes to generate information about the yield of a crop over time and/or during a growing season.

As noted earlier, this not simply an academic exercise. This type of analysis is envisioned for various points in the growing season and would be accessed by individual farmers to determine the direction for their specific crops as well as for others in the government or in private investment industry that wish to determine the future prospects for a crop in question. Thus the economic potential for the present invention is significant.

The present invention is implemented on a Sun Microsystem server which runs the data base, analysis algorithms, server software and four connecting to a work stations. The work stations are connected over a network which may be a local area network (LAN), a wide area network (WAN), and/or work stations connected to the Internet. These network connections are but one example of the type of network connection and are not meant as a limitation. For example, workstations may be connected in a wireless fashion to the server of the present invention simply by means of a transceiver located at the workstation location and at the server. Alternatively connections exist between remote work stations operating wirelessly to Internet service providers and thence to the Internet for connection to the server of the present invention.

Analysis algorithms that are stored in the server and used for the present invention are, for example, and without limitation, multi-variate statistics and artificial intelligence algorithms such as clustering algorithms, multi-variate factor analysis, principal axes factor analysis, and other types of multi-variate algorithms which are capable of being exercised by the server of the present invention. In addition curve fitting algorithms of various types known in the art are stored and available to the analyst as are various patent recognition algorithms, flux analysis together with metabolic control analysis and various visualization options for display of data. This information is integrated together with proteomics and gene expression data bases to allow correlations with these types of data as well.

It is also the situation that measurement of certain of the response variables that are of an historical nature and those of a laboratory nature follow different protocols for their collection. For example, collecting information on the sugar content of grapes may take place at one frequency during the course of the growing season while information on genetic expression may take place at an entirely different frequency during the growing season for the same crop. Conversely, crops that are harvested only once every tens of years, such as wood and paper product type crops may have sampling frequencies that are radically different than those crops that are grown and harvested within a single season.

**Brief Description Of The Drawings**

Figure 1 illustrates response variable types at a particular time.

Figure 2 illustrates relationships between response variable and collection protocols.

Figure 3 illustrates the conceptual database structure.

Figure 4 illustrates the overall system architecture.

**Detailed Description Of The Invention**

Referring to Figure 1, the conceptual framework for response variables is illustrated. Response variables 100 comprise environmental data 102, laboratory data 104, metabolic data 106, and genetic data 108. Environmental data 102 may comprise several types of environmental data 110, 112, and 114 which may be rainfall, humidity, temperature, and indeed any other type of environmental information that may be important to the growth cycle of a particular species under analysis.

Laboratory data 104 comprises multiple types of data as well, herein illustrated as type 1 116, type 2 118, type 3 120, and type 4 122.

Various types of metabolites are also recorded in metabolic data 106. Here metabolites 124, 126, 128, and 130 are recorded with respect to their presence as well as their concentrations.

Finally genetic data 108 is also recorded as a series of observations with the presence of various mutant genes for a particular species. Thus, genetic data 132, 134, 136, and 138 are observed and their presence recorded.

It should be noted that all of these response variables 100 are recorded at a particular time during the growing cycle T1.

Referring to Figure 2, collection information in the form of data protocols is conceptually illustrated. Response variable data 100, which comprises response variables 102, 104, and 108 are all associated with a data protocol information 160 which comprises data protocols 162, 164, and 166. These data protocols are each associated with the response variables so that an analyst/researcher can determine how a particular response variable was derived and what the various data sampling schemes were that were associated with each particular response variable.

Referring to Figure 3, the conceptual database structure is illustrated. Response variables 100 are recorded for a particular time T1 during the growing season (GS1), 180. The same response variables are also recorded for times T2 170 and TN 172 which represent various times for sampling during growing season number 1, 180. This same type of response variable sampling during different times in a growing season is also recorded for growing season 2, 182 and growing season N 184.

All of this information regarding response variables 100 recorded at different times for different growing seasons are all recorded for a particular species 190. This sampling and recording of data is also done for additional species 192 and 194. Thus a particular response variable may be analyzed across species, across growing seasons, and across different sampling times during a particular growing season. All of this information may be parsed and analyzed in a multi-variate way. It should be noted that the specific number of growing seasons, species, response variables, and sampling times are all illustrative in nature and are not meant to be limiting. In practice, there will be many sampling times during a particular growing season, and many species of commercial value may be analyzed in the fashion noted in the present invention.

Referring to Figure 4, the overall system architecture is illustrated. Metabolic processor 200 has a series of supporting databases. Raw data from various environmental, laboratory, and other measurements are stored for the data types and sampling first noted in the conceptual database structure (Figure 3). These data are called upon for the various analysis desired by the scientists.

A protocol database 204 is also stored and is related to the various data stored in the raw database 202. In this fashion, an analyst can analyze any single piece of data and determine the protocol that was used to obtain it.

Types of response variables are stored in a response variable database 206 which allows an analyst to determine what types of data may be resonant in the database and what types of data could be obtained in order to support any analysis task.

An algorithmic database 208 is available to the analyst for subsequent loading on the metabolic processor 200. This provides the analyst with a wide variety of multi-variate analyses such as, and without limitation, clustering algorithms, flux analysis, metabolic control analysis, multi-variate factor analysis, principle axes factor analysis, curve fitting, pattern-recognition, and similar tools. All of these tools are stored in algorithmic database 208 and can be loaded on metabolic processor 200 to serve as the basis for analyzing raw data 202 concerning any particular species or trends of data within the species.

Once a particular algorithm in combination with certain response variables and raw data are determined to be useful, that specific analysis is stored in an analysis database 220. By storing the appropriate analysis steps, a subsequent scientist can access metabolic processor 200 via workstations 210, 212 and request a specific analysis for a specific species. This algorithm will be then retrieved from the analysis database 220 which will automatically cause the appropriate raw data to be retrieved and analysis results to be output to the researcher at the workstation.

A series of workstations are connected to the metabolic processor 200. As illustrated, workstations 210 and 212 are connected via a local area network to the metabolic processor 200.

Metabolic processor 200 can also be accessed over a network 214 by remote workstations 216 and 218. Examples of such a network can be an intranet, the internet, or any other network suitable for providing remote access to a central processor.

As noted earlier, the present system is implemented on a sun microsystems server for running the database, analysis algorithms, and server software. This will allow any computer with the web browser to act as a client for the metabolic processor 200.

Generally any type of workstation such as an IBM PC or compatible running, for example, a Pentium processor having local storage, and output capability will be suitable for a client station for the system.

Various technologies will serve as the basis for collecting raw data of a laboratory nature concerning species of interest. For example, and without limitation, gas chromatography, mass spectrometry, HPLC, LCMS, TLC, UV-VIS, SWIR, MWIR, LWIR, raman spectrometry, and various bio-sensor information can all be collected and tagged with the appropriate protocol used for the collection and associated with a particular species and timed during the growing season during which the samples were taken.

Using the metabolic information recorded in the database structure, and using the system of the present invention, qualitative studies can be accomplished to determine which metabolites are expressed and potentially discover novel compounds which are indicative of the quality of a particular harvest.

In addition, quantitative analysis can be conducted to measure concentrations of metabolites during the course of the growing season in order to determine a trend for the particular crop in question. In this fashion, it will ultimately be possible to create predictive models to assist in optimizing any particular crop for desired characteristics of yield and quality. Further, by bringing together these disparate data types, it will be possible for a scientist or analyst to be able to evaluate in a streamline fashion data that heretofore, has not been able to be combined in any meaningful fashion within a database.

With knowledge of how to conduct a specific type of study, an analyst can simply input a desired end product for an analysis, such as, for example and without limitation, how much orange juice can a particular crop produce? The system can then select the appropriate analysis algorithm, pull the data from the database, and create the predictive model or response. If, on the other hand, vitamin or supplement products are desired from the orange crop, a different model may be run, using perhaps a different predictive algorithm from the database.

In the alternative, a grower can ask the system what type of data and sampling rates are required if that grower is to make a prediction for an optimized amount of product from a given crop.

A system and method for metabolic profiling has been described. It will be apparent to those skilled in the art that other types of data can be brought into the system for analysis, the types of analysis tools can be stored in the analysis database for use by scientists, other types of protocols for obtaining different types of data may also be created and stored for later access by the scientist without departing from the scope of the invention as disclosed.

I claim:

1. An apparatus for metabolic profiling comprising:

a processor;

a database connected to the processor for storing data concerning a plurality of plant species;

at least one workstation for accessing the processor; and

a database comprising metabolic data concerning the plurality of plant species.

2. The system for metabolic profiling of claim 1 wherein the database further comprises an algorithmic database comprising computer programs for multi-variate analysis of the metabolic data.

3. The system for metabolic profiling of claim 2 wherein the computer programs for multi-variate analysis are taken from the group consisting of clustering algorithms, flux analysis, metabolic control analysis, system identification, multi-variate factor analysis, principle axes factor analysis, pattern-recognition, and curve fitting programs.

4. The system of metabolic profiling of claim 1 when the database further comprises protocol data correlating metabolic data and how metabolic data was obtained.

5. The system for metabolic profiling of claim 1 wherein the database further comprises workflow routines for established analysis techniques.

6. The system for metabolic profiling according to claim 1 wherein the database further comprises environmental data.

7. The system for metabolic profiling according to claim 1 wherein the database further comprises species developmental data.

8. The system for metabolic profiling according to claim 6 wherein the processor further comprises instructions for analyzing the same species as a function of the environmental data.

9. The system for metabolic profiling according to claim 7 wherein the processor further comprises instructions for analyzing the same species as a function of the developmental data.

10. The system for metabolic profiling according to claim 6 wherein the processor further comprises instructions for analyzing different species as a function of the environmental data.

11. The system for metabolic profiling according to claim 1 wherein the metabolic data comprises data from analysis of plant species using laboratory sensors.

12. The system for metabolic profiling according to claim 11 wherein the laboratory techniques comprise techniques from the group consisting of GS-MS, HPLC, LCMS, TLC, UV-VIS, SWIR, MWIR, LWIR, raman spectrometry, and biological sensing.

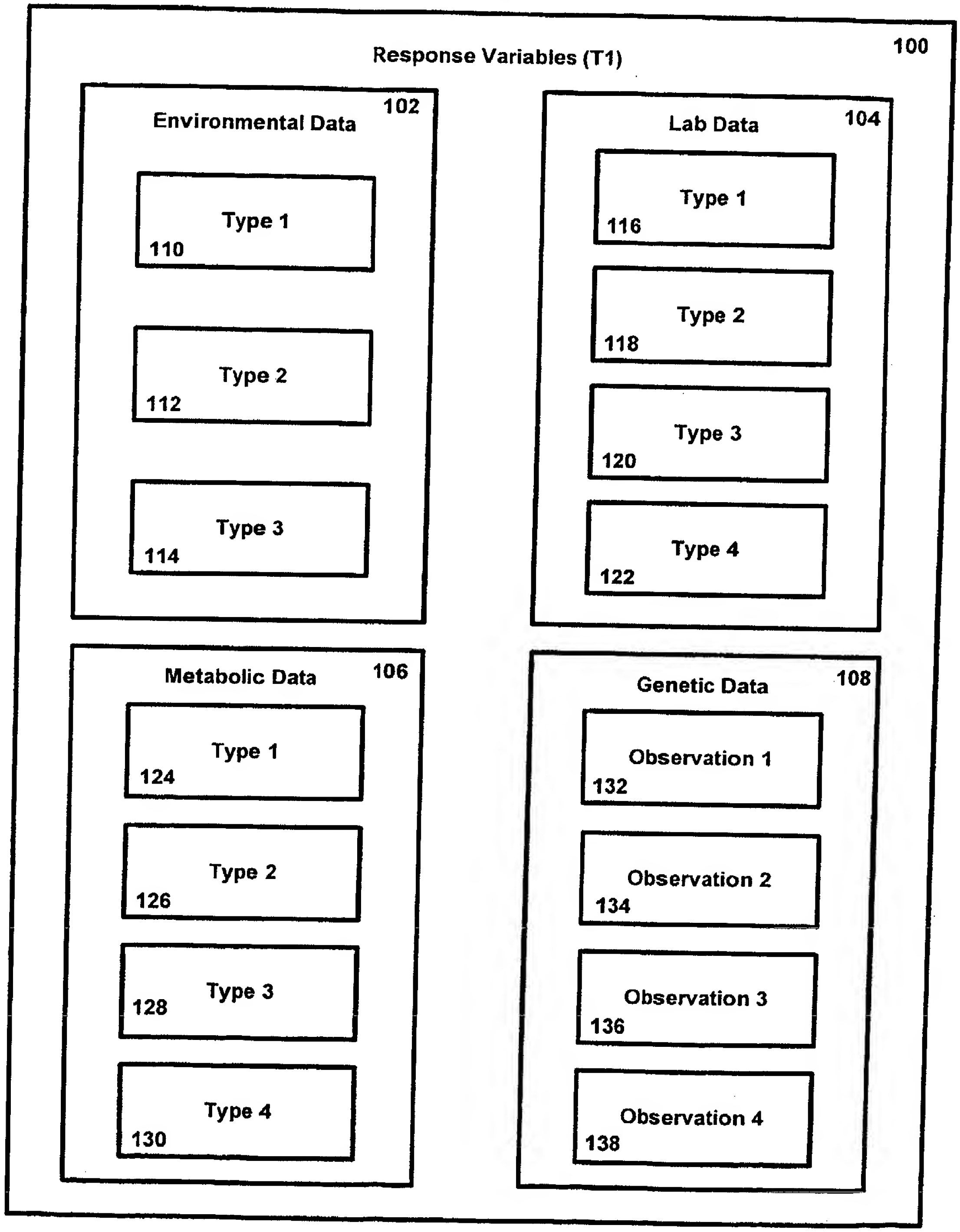

Response Variables (T1)
100
Environmental Data
102
Type 1
110
Type 2
112
Type 3
114
Lab Data
104
Type 1
116
Type 2
118
Type 3
120
Type 4
122
Metabolic Data
106
Type 1
124
Type 2
126
Type 3
128
Type 4
130
Genetic Data
108
Observation 1
132
Observation 2
134
Observation 3
136
Observation 4
138

Figure 1

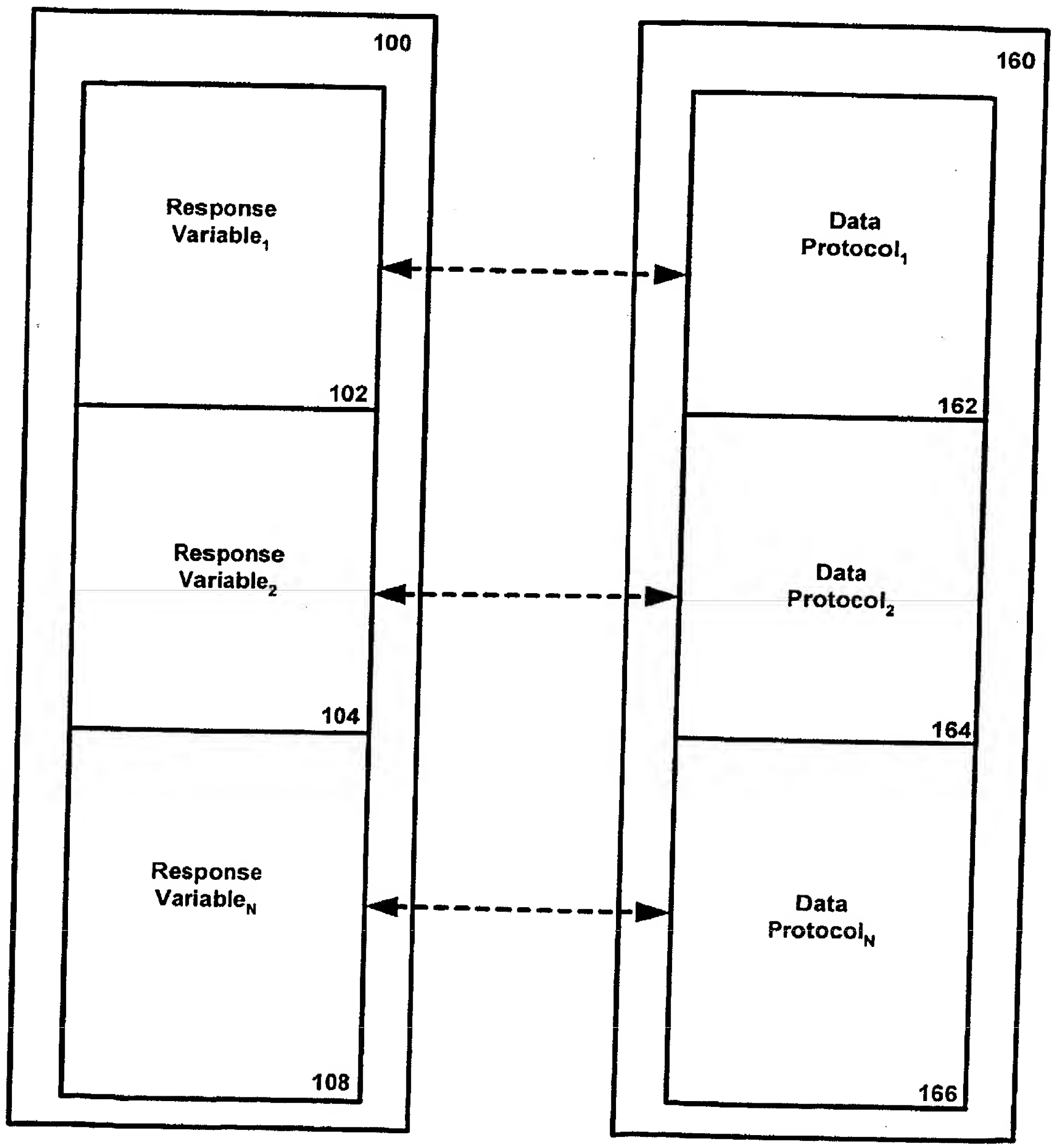

100
Response Variable$_1$
102
Response Variable$_2$
104
Response Variable$_N$
108
160
Data Protocol$_1$
162
Data Protocol$_2$
164
Data Protocol$_N$
166

Figure 2

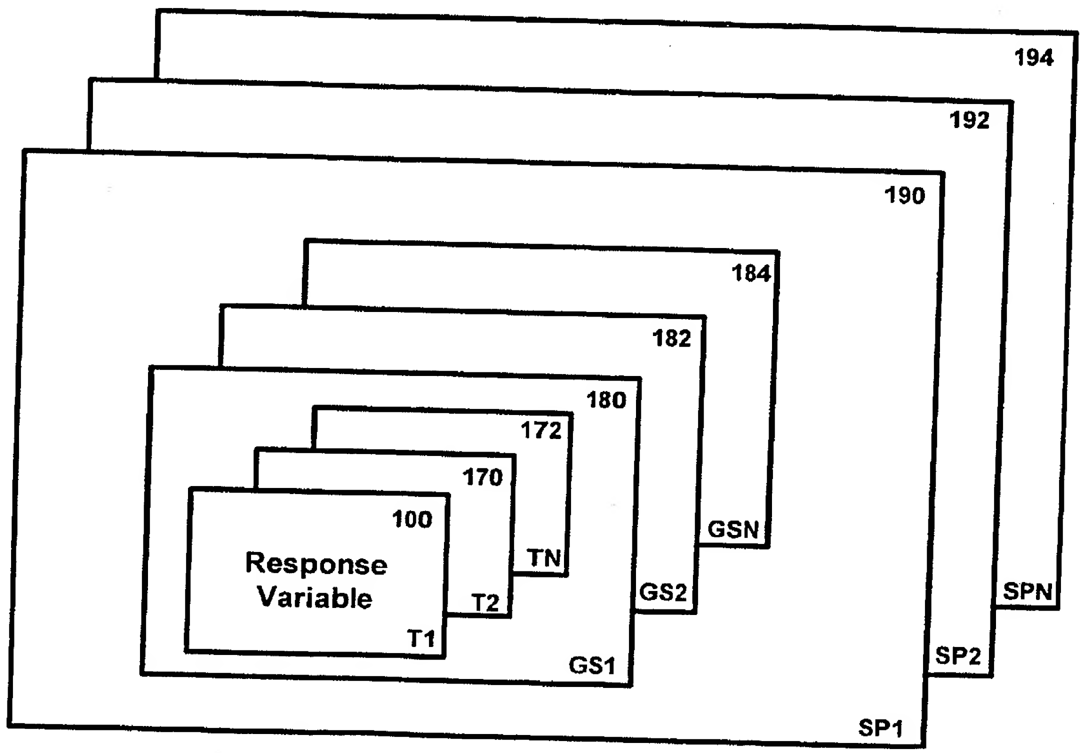

194
192
190
184
182
180
172
170
100
Response Variable
T1
T2
TN
GS1
GS2
GSN
SP1
SP2
SPN

Figure 3

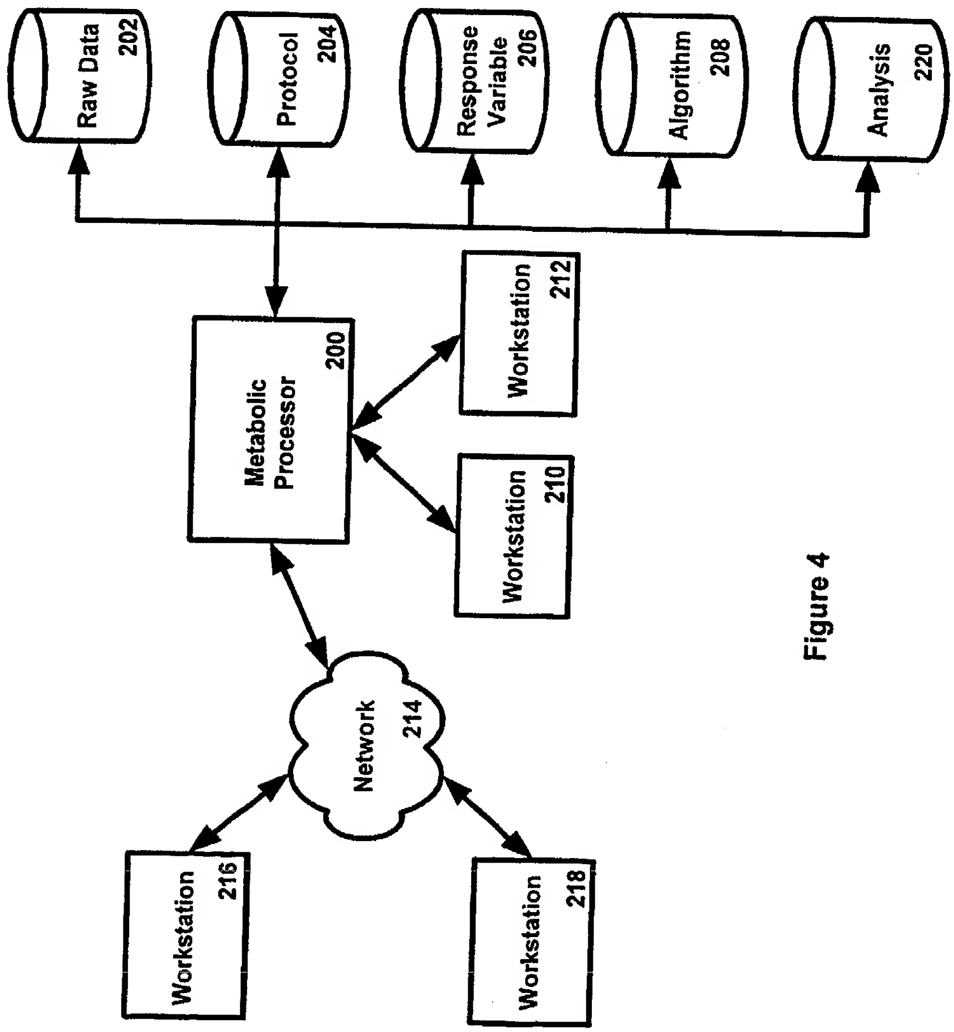

Raw Data
202
Protocol
204
Response
Variable
206
Algorithm
208
Analysis
220
Metabolic
Processor
200
Workstation
212
Workstation
210
Network
214
Workstation
216
Workstation
218

Figure 4

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 June 2001 (07.06.2001)

PCT

(10) International Publication Number
WO 01/40896 A3

(51) International Patent Classification[7]: G06F 19/00

(21) International Application Number: PCT/US00/33069

(22) International Filing Date: 6 December 2000 (06.12.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/454,991 6 December 1999 (06.12.1999) US

(71) Applicant: NATIONAL CENTER FOR GENOME RESOURCES [US/US]; 2935 Rodeo Park Drive East, Santa Fe, NM 87505 (US).

(72) Inventor: MENDES, Pedro; Apt. 512, 2500 Sawmill Road, Santa Fe, NM 87505 (US).

(74) Agents: ROBERTS, Jon, L. et al.; Roberts Abokhair & Mardula, LLC, Suite 1000, 11800 Sunrise Valley Drive, Reston, VA 20191 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

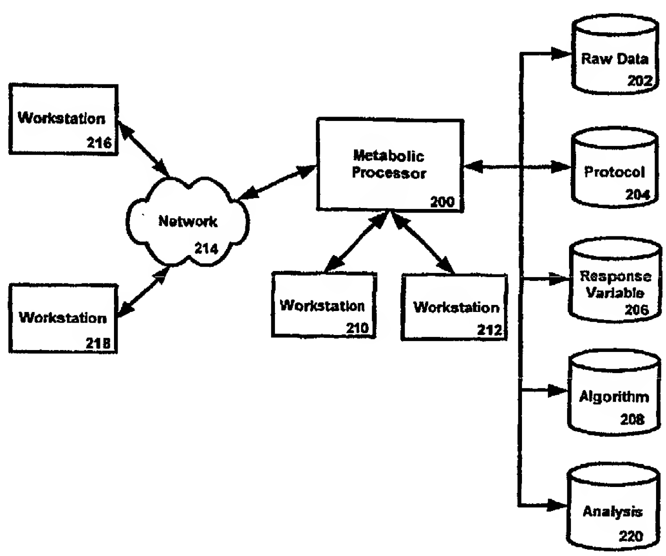
(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**
— *with international search report*

(88) Date of publication of the international search report:
21 March 2002

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: SYSTEM AND METHOD FOR METABOLIC PROFILING




(57) Abstract: A system and method for metabolic profiling for different species. The system comprises data from a wide variety of laboratory sensors and other measurements combined with an algorithmic database allowing multi-variate analysis of the data to be performed. Protocol information is stored concerning the data resident within the database. Data is retrieved throughout the growing season, across growing seasons, and for a wide variety of response variables, and for different species, thus allowing a wide variety of analysis across species and within species to take place. Data is stored concerning the workflow of any particular analysis so that scientists can build upon successful analyses to create further novel methods of analyzing metabolic data.

WO 01/40896 A3

# INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 00/33069

**A. CLASSIFICATION OF SUBJECT MATTER**
IPC 7 G06F19/00

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, PAJ, WPI Data

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category ° | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| E | WO 01 09711 A (PARADIGM GENETICS INC ;DAVIS KEITH R (US); BOYES DOUGLAS C (US); G) 8 February 2001 (2001-02-08) abstract; claims 1-7 | 1-12 |
| E | WO 01 16857 A (REALTIMEHEALTH COM INC) 8 March 2001 (2001-03-08) abstract | 1-12 |

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 20 November 2001 | 27/11/2001 |

| Name and mailing address of the ISA | Authorized officer |
|---|---|
| European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016 | Filloy García, E |

1

Form PCT/ISA/210 (second sheet) (July 1992)

# INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 00/33069

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category ° | Citation of document, with indication,where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| A | FOTEDAR L K ET AL: "PLANT CONDITION REMOTE MONITORING TECHNIQUE" IGARSS 1996. INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM: REMOTE SENSING FOR A SUSTAINABLE FUTURE. LINCOLN, NE, MAY 28 - 31, 1996, INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM. IGARSS, NEW YORK, IEEE, US, vol. 1, 28 May 1996 (1996-05-28), pages 239-242, XP000659599 ISBN: 0-7803-3069-2 the whole document --- | 1-12 |
| A | WO 99 21118 A (DOLEZAL WILLIAM E ;PIONEER HI BRED INT (US); BERRY JAMES A (US); S) 29 April 1999 (1999-04-29) abstract ----- | 1-12 |

1

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 00/33069

| Patent document cited in search report | | Publication date | Patent family member(s) | | | Publication date |
|---|---|---|---|---|---|---|
| WO 0109711 | A | 08-02-2001 | AU | 6505600 | A | 19-02-2001 |
| | | | WO | 0109711 | A2 | 08-02-2001 |
| WO 0116857 | A | 08-03-2001 | AU | 6939700 | A | 26-03-2001 |
| | | | AU | 7077500 | A | 26-03-2001 |
| | | | AU | 7826100 | A | 26-03-2001 |
| | | | WO | 0116857 | A2 | 08-03-2001 |
| | | | WO | 0116858 | A2 | 08-03-2001 |
| | | | WO | 0116859 | A2 | 08-03-2001 |
| | | | US | 2001037340 | A1 | 01-11-2001 |
| WO 9921118 | A | 29-04-1999 | US | 6014451 | A | 11-01-2000 |
| | | | AU | 9806898 | A | 10-05-1999 |
| | | | BR | 9813086 | A | 22-08-2000 |
| | | | EP | 1025529 | A1 | 09-08-2000 |
| | | | WO | 9921118 | A1 | 29-04-1999 |